

THREE-DIMENSIONAL IMAGE GENERATING DEVICE

Patent Number: JP10255070
Publication date: 1998-09-25
Inventor(s): OTSUKA NAOHIRO; OTANI ATSUSHI
Applicant(s): ATR CHINOU EIZO TSUSHIN KENKYUSHO:KK
Requested Patent: ☐ JP10255070
Application Number: JP19970056285 19970311
Priority Number(s):
IPC Classification: G06T15/70
EC Classification:
Equivalents: JP3055666B2

Abstract

PROBLEM TO BE SOLVED: To easily generate an interface using a required face by generating a highly realistic expression following a change in the expression of an inputted moving image of a certain person on a three-dimensional face wire frame model (3D face WFM) corresponding to the face of another person.

SOLUTION: An expression recognition part 1 recognizes the expression of a face input moving image by calculating state probability to be a 1st parameter based on a feature vector obtained by measuring the variation of face elements of a certain person. A 3D face WFM generation part 9 generates a 3D face WFM of a 2nd person. An expression conversion part 3 converts the state probability of the recognition part 1 into a face model deformation parameter to be a 2nd parameter for deforming the 3D face WFM, based on the variation. An expression generation part 5 generates the 3D face image of the 2nd person by deforming the 3D face WFM generated by the generation part 9, based on the face model deformation parameter calculated by the conversion part 3.

Data supplied from the esp@cenet database - I2

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-255070

(43) 公開日 平成10年(1998) 9月25日

(51) Int.Cl.⁹
G 0 6 T 15/70

識別記号

F I
G 0 6 F 15/62

3 4 0 K

審査請求 有 請求項の数 1 O L (全 5 頁)

(21) 出願番号 特願平9-56285

(22) 出願日 平成9年(1997) 3月11日

(71) 出願人 595147700

株式会社エイ・ティ・アール知能映像通信
研究所
京都府相楽郡精華町大字乾谷小字三平谷5
番地

(72) 発明者 大塚 尚宏

京都府相楽郡精華町大字乾谷小字三平谷5
番地 株式会社エイ・ティ・アール知能映
像通信研究所内

(72) 発明者 大谷 淳

京都府相楽郡精華町大字乾谷小字三平谷5
番地 株式会社エイ・ティ・アール知能映
像通信研究所内

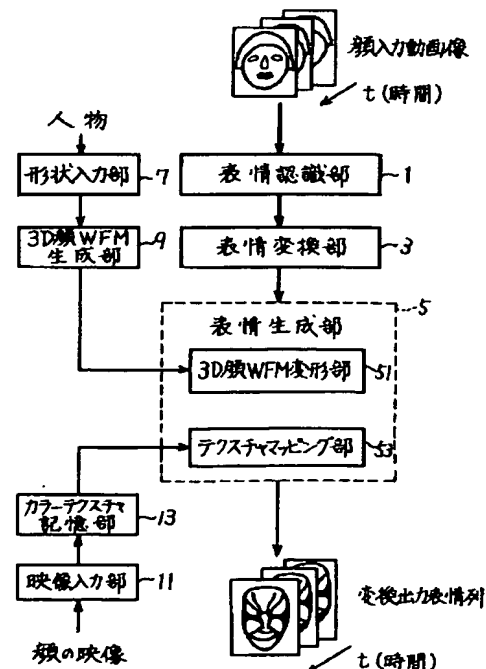
(74) 代理人 弁理士 深見 久郎 (外2名)

(54) 【発明の名称】 3次元画像作成装置

(57) 【要約】

【課題】 ある人物の表情を他の人物の表情に変換した3次元顔画像を作成する3次元画像作成装置を提供する。

【解決手段】 入力される第1の人物の顔動画像において顔要素の変形量を計測し、その変形量から得られる特徴ベクトルに基づいて状態確率を求めることにより、上記第1の人物の表情を認識する表情認識部1と、第2の人物の3次元顔モデルを生成する3D顔WFM生成部9と、上記状態確率から表情変換行列を用いてモデル変形量を算出する表情変換部3と、上記モデル変形量に基づいて上記3次元顔モデルを変形し第2の人物の3次元顔画像を生成する表情生成部5とを備える。



【特許請求の範囲】

【請求項1】 入力される第1の人物の顔動画像において、顔要素の変形量を計測し、前記変形量から得られる特徴ベクトルに基づいて第1のパラメータを決定することにより前記第1の人物の表情を認識する表情認識手段と、
第2の人物の3次元顔モデルを作成するモデル作成手段と、
前記第1のパラメータを、前記第2の人物の表情変化から得られる顔要素の変化率に基づいて、前記第2の人物の3次元顔モデルを変形するための第2のパラメータに変換する表情変換手段と、
前記第2のパラメータに応じて前記3次元顔モデルを変形し、前記第2の人物の3次元顔画像を作成する表情生成手段とを備える、3次元画像作成装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、3次元画像を作成する3次元画像作成装置に関するものである。

【0002】

【従来の技術】異なる場所の人物同士の仮想的なシーンを介した新しいコミュニケーション手段などの仮想的環境の実現のためには、3次元人物像の表示法が重要であり、実際の人物の姿の忠実な再現、人物の姿を別の姿に変換する仮想変身、の大別して2通りの表示法が存在すると考えられる。ここで、仮想変身システムの応用としては娯楽や芸術がまず考えられるが、一方でテレビ電話などの映像通信が必ずしも盛んでない要因の1つに、自分の素顔が相手のテレビに映ることへの抵抗感があることを考えると、仮想変身は映像通信の普及のためのポイントの1つということが出来る。

【0003】ある人物の表情を別の人物の表情に変換する従来の技術としては、顔の3次元コンピュータグラフィック（CG）モデルにおいて、ある人物より検出された表情を再現する技術がある。

【0004】

【発明が解決しようとする課題】しかしながら、これら従来の表情再現法に関する技術では、顔の表情筋の動作に伴う皮膚表面の収縮量の推定結果に基づき、3次元ワイヤフレームモデル（3DWFM）を変形して表情を生成するにとどまっていた。すなわち、これらの従来技術では、表情認識における特徴抽出処理と、表情生成とが概念的には直結されていたため、異なる人物同士などの表情変換処理は全く考慮されていなかった。

【0005】そこで、本発明は、表情筋の動作量の検出結果だけでなく、表情の認識処理結果も併せて利用することにより、ある人物の表情を他の人物の表情に変換して3次元顔画像を生成する3次元画像作成装置を提供することを目的とする。

【0006】

【課題を解決するための手段】請求項1に係る3次元画像作成装置は、入力される第1の人物の顔動画像において、顔要素の変形量を計測し、変形量から得られる特徴ベクトルに基づいて第1のパラメータを決定することにより第1の人物の表情を認識する表情認識手段と、第2の人物の3次元顔モデルを作成するモデル作成手段と、第1のパラメータを、第2の人物の表情変化から得られる顔要素の変化率に基づいて、第2の人物の3次元顔モデルを変形するための第2のパラメータに変換する表情変換手段と、第2のパラメータに応じて3次元顔モデルを変形し、第2の人物の3次元顔画像を作成する表情生成手段とを備えるものである。

【0007】

【発明の実施の形態】以下、本発明の実施の形態を図面を参照して詳しく説明する。

【0008】図1は、本発明の実施の形態に係る3次元画像作成装置の構成を示すブロック図である。図1に示されるように、この3次元画像作成装置は、ある人物の顔入力動画像において、顔要素の変形量を計測し、その変形量から得られる特徴ベクトルに基づいて、第1のパラメータとしての状態確率を算出することにより上記顔入力動画像の表情を認識する表情認識部1と、他の人物の顔の形状を入力する形状入力部7と、形状入力部7に入力されたデータに基づいて3次元（3D）顔ワイヤフレームモデル（WFM）を生成する3D顔WFM生成部9と、表情認識部1で算出された状態確率を、表情認識部1で計測された変形量に基づいて、上記3D顔WFMを変形するための第2のパラメータとしての顔モデル変形パラメータに変換する表情変換部3と、上記他の人物の顔の映像を入力する映像入力部11と、映像入力部11に入力された映像に基づいて顔表面の色彩情報を獲得するカラーテクスチャ記憶部13と、表情生成部5とを備える。ここで、表情生成部5は、表情変換部3で算出された顔モデル変形パラメータに基づいて3D顔WFM生成部9で生成された3D顔WFMを変形する3D顔WFM変形部51と、変形された3D顔WFMにカラーテクスチャ記憶部13で得られた色彩情報に基づいて色彩を付するテクスチャマッピング部53とを含む。

【0009】以下において、上記の主要な各部について詳しく説明する。図2は、図1に示される表情認識部1の動作を示すフローチャートである。図2に示されるように、まず最初にステップS1では、入力される動画像のフレーム毎に顔の各部位の速度ベクトルを推定する。具体的には、顔入力動画像の隣接する2つのフレームからオブティカルフローアルゴリズムを用いて速度ベクトルを求める。画像処理においては、たとえば、右目および口をそれぞれ中心とする2つの領域に対して行なう。この2つの領域の位置合わせは、動画像の第1フレームにおいて手動で行なう。なお、動画像はヘルメットに装着された小型CCDカメラにより撮影されているので頭

部の動きの影響はない。

【0010】次に、ステップS2では、ステップS1で得られた速度ベクトルの各成分に離散フーリエ変換を施し、フーリエ変換係数の低周波成分（右目・口領域から7、8個、合計15個）を特徴ベクトルとして抽出する。これらの特徴ベクトルの成分は、表情変化の開始時および終了時に値が0となり2つの時刻での区別がつかないため、右目および口領域の速度ベクトルの自乗和を時間積分した値を特徴ベクトルの成分として加える。

【0011】次に、ステップS3、S4では、それぞれ隠れマルコフモデル（HMM）を用いて各表情カテゴリの出力・状態確率を算出する。

【0012】HMMは状態数Nと、状態 S_i から次の時刻に状態 S_j に移移する確率 a_{ij} 、状態 S_j への移移に

$$b_j(V) = \sum_{k=1}^M c_{jk} N[V, \mu_{jk}, U_{jk}] \quad \dots(1)$$

【0015】ここで、j番目の混合要素におけるパラメータ μ_{jk} 、 U_{jk} は学習データからクラスタリング手法により初期値を設定し、Baum-Welchアルゴリズムを用いて求める。

【0016】一方、ステップS4での状態確率算出においては、移移確率 a_{ij} は乱数により初期値を設定し、Baum-Welchアルゴリズムを用いて求められる。

$$p_i(t) = \left[\sum_{j=1}^N p_j(t-1) a_{ji} \right] b_i(V_t) \quad \dots(2)$$

【0019】次に、ステップS5では、筋肉の収縮が終了した状態 S_3 の確率が最大となる表情カテゴリを認識結果とするが、表情変換においては、筋肉が収縮中の状態 S_2 の確率を表情変換部3に出力する。

【0020】表情変換部3では、次式のように表情認識部1で得られた表情カテゴリの確率（状態 S_2 の確率）

$$D = TP \quad \dots(3)$$

【0022】表情変換行列Tの(i, j)要素は各表情に対する変形パラメータ D_i のフレーム間隔当りの変化量を表わす。この行列Tの要素の設定には、Ekmanらが表情分析の手法として提案したFACS（Facial Action Coding System）を適用する。

【0023】具体的には、行列Tの要素の値は、表情と顔モデル変形パラメータとの相関がある要素に対して、上記他の人物の表情変化を示す顔動画像から予め得られる筋肉の収縮速度、顎の回転の角速度、目および口の形状変化の高さあるいは幅の変化率であり、変換した表情を形成する他の人物の顔モデルの種類に依存しないものである。

【0024】表情生成部5は、3D顔WFM変形部51で29個の顔モデル変形パラメータにより3D顔WFMを変形し、テクスチャマッピング部53でその3D顔WFMの上にテクスチャをマッピングする。顔モデル変形

に伴って特徴ベクトルVを出力する出力確率 $b_j(V)$ によって特徴付けられる。ここでは、表情の変化が3つの筋肉の状態（弛緩、収縮中、収縮終了）の変化により引き起こされると仮定し、さらに、その変化は1方向であるとして図3に示すLeft-to-right型のHMMによりモデル化する。

【0013】ステップS3で算出される出力確率 $b_j(V)$ は、以下のように平均値ベクトル μ 、共分散行列Uをパラメータとする多変数正規分布 $N[V, \mu, U]$ の荷重平均（混合数M、荷重 c_{jk} ）により近似する。

【0014】

【数1】

【0017】特徴ベクトルVの時系列 V_t が与えられたとき、HMMの状態 S_i の時刻tでの第1のパラメータとしての状態確率 $p_i(t)$ は以下の前向きパスアルゴリズムにより求められる。

【0018】

【数2】

を要素とするベクトルPに表情変換行列Tを乗算して、以下で説明する第2のパラメータとしての顔モデル変形パラメータを要素とするベクトルDを求める。

【0021】

【数3】

...

パラメータは、具体的には、表情筋の収縮につき自由度26、顎の回転につき自由度1、目の開閉につき自由度1、口を尖らすことにつき自由度1を有するものである。表情筋の中で、大頬骨筋などの線型筋肉は骨（筋肉の始点と呼ぶ）と皮膚（筋肉の終点と呼ぶ）とに端点を持ち、筋肉の収縮により終点の近傍の皮膚を始点の方向に移動させる作用をする。ここで、この作用を3D顔WFMにおいては終点の近傍の頂点を始点の方向に移動させることにより実現する方法も考えられているが、本実施の形態においては、皮膚の弾性体モデルに基づいて皮膚の移動量を算出する。

【0025】まず、1次元のモデルで考える。x軸の原点に筋肉の終点があり、x軸の負方向に筋肉が収縮したときの点xでのずれを $u(x)$ とする。皮膚の微小部分に働く両側からの応力と皮膚の位置のずれに伴う張力とは釣り合っているため、応力の差 F_1 および張力 F_2 を

表わす以下の式を等しいとおいて微分方程式を解くと解
(6)が得られる。

【0026】

【数4】

$$F_1 = YA \frac{u(x + \Delta x) - u(x)}{\Delta x} \xrightarrow{\Delta x \rightarrow 0} YAu'(x) \quad \dots(4)$$

【0027】

【数5】

$$F_2 = ku(x) \quad \dots(5)$$

【0028】

【数6】

$$u(x) = u_0 \exp\left(-\frac{k|x|}{YA}\right) \quad \dots(6)$$

【0029】ここで、Y、k、Aは皮膚のヤング率、張力係数、厚みである。以上のモデルを2次元に拡張して原点に筋肉の終点があり、始点が負のx軸上にある座標系で考えると、筋肉がx軸方向に収縮したときの点

(x, y)でのx軸方向のずれu(x, y)は以下の式で書ける。

【0030】

【数7】

$$u(x, y) = u_0 \exp\left(-\frac{k|x|}{YA} - \frac{k|y|}{SA}\right) \quad \dots(7)$$

【0031】ここで、Sは皮膚の剛性率である。剛性率の値をヤング率の1/3とし、係数を新たにαと定義すると式(6)は次式のように書ける。

【0032】

【数8】

$$u(x, y) = u_0 \exp\{-\alpha(|x|+3|y|)\} \quad \dots(8)$$

【0033】ここで、筋肉の幅を考慮して式(7)のyとして筋肉の端からの長さが用いられる。

【0034】係数αは皮膚の柔軟性に比例しており、値が大きいくほど柔軟となり筋肉の影響は局所化される。ここで、係数αを操作することにより、表情を強調したり抑制したりすることができる。

【0035】たとえば、表情の生成において、顎の回転は、まず顎骨の近傍の頂点を顎の回転角分回転させ、次に、隣り合う頂点の移動量の変化が滑らかになるように調整する。そして、口の輪郭部は式(6)を用いて張力によるずれを求めて補正する。

【0036】一方、眼輪筋と口輪筋は、楕円型の形状で楕円の中心に向かって収縮するため、線型筋のように簡単にモデル化することはできない。そこで、目の開閉および口を尖らす動作に対して3Dモデルの頂点の速度ベクトルを予め用意しておき、動作の大きさに応じて速度ベクトルの長さを変えることにより表情を生成する。

【0037】以上より、本発明の実施の形態に係る3次元画像作成装置によれば、ある人物の入力動画像に対して、入力動画像の表情の変化に追従したリアリズムの高い表情を他の人物の顔に対応する3D顔WFM上に生成することができ、ユーザの所望の顔を用いたインタフェースを容易に作成することができる。また、ある人物の

表情を強調等し、その強調等された表情を他の人物の表情として表わすことができるため、ある人物の心情を容易に理解させることのできる3次元顔画像を作成することができる。

【0038】

【発明の効果】請求項1に係る3次元画像作成装置によれば、動画像中のある人物の表情を他の人物の3次元顔モデルにおける表情に変換した3次元顔画像を作成することができる。

【図面の簡単な説明】

【図1】本発明の実施の形態に係る3次元画像作成装置の構成を示すブロック図である。

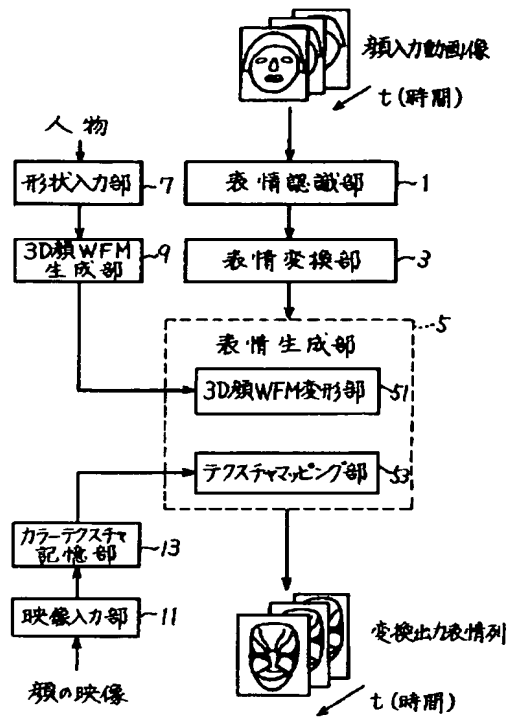
【図2】図1に示される表情認識部の動作を説明するためのフローチャートである。

【図3】状態数が3のときのLeft-to-right型HMMの構成を示す図である。

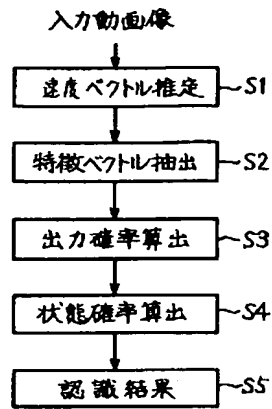
【符号の説明】

- 1 表情認識部
- 3 表情変換部
- 5 表情生成部
- 9 3D顔WFM生成部
- 51 3D顔WFM変形部

【図1】



【図2】



【図3】

